Integrating Multimedia Archives: The Architecture and the Content Layer

Manolis Wallace, *Member, IEEE*, Thanos Athanasiadis, *Member, IEEE*, Yannis Avrithis, *Member, IEEE*, Anastasios Delopoulos, *Member, IEEE*, Stefanos Kollias, *Member, IEEE*

***Abstract*— During the last few years numerous multimedia archives have made extensive use of digitized storage and annotation technologies. Still, the development of single points of access, providing common and uniform access to their data, despite the efforts and accomplishments of standardization organizations, has remained an open issue, as it involves the integration of various large scale heterogeneous and hetero-lingual systems. In this paper, we describe a mediator system that achieves architectural integration through an extended 3-tier architecture and content integration through semantic modeling. The described system has successfully integrated five multimedia archives, quite different in nature and content from each other, while also providing for easy and scalable inclusion of more archives in the future.**

***Index Terms*— multimedia archives, architectural integration, mediator, 3-tier architecture, message oriented middleware, semantic modeling, concept taxonomy, document analysis, MPEG-7.**

I. Introduction

During the last decade, the cost of storage and wide area communication services has decreased, while their capacity increased dramatically. This fact, along with the increasing penetration of e-commerce applications, has made *digital storage, annotation and access* of multimedia information a mature and viable choice for content holding organizations and individuals. Numerous multimedia archives have, either totally or incrementally, turned to the utilization of digitized archive technologies. The content of these archives can be made accessible, depending upon copyright, policy and security decisions, over the Internet in a cost efficient, time efficient, anyplace, anytime fashion.

On the other hand, one of the main problems of multimedia archiving has also been inherited to their digital descendants. For traditional archives, where raw media were stored in the form of analog hard copies, search was not an easy task as a human had to either go through a separate annotation archive, or, ideally, search using keywords in a custom, proprietary, metadata database. Much similarly to the case of books in a library that have not been indexed, information stored in a multimedia archive that cannot be searched, identified and accessed easily is practically unavailable.

In order to provide for more *efficient search services* in the always augmenting space of available digital multimedia content, several systems have been proposed and several research projects and initiatives have been funded, making important contributions to theoretical fields ranging from multimedia signal processing, computer vision, multimedia database and knowledge management to artificial intelligence, human computer interaction and information retrieval. Still, considering the number and diversity of multimedia archives existing worldwide, being able to search in each one of them independently but not in all of them at the same time through a common interface is much like having independent indices for each corridor in a library. When the library becomes larger, data is once again as good as non existing.

Current and evolving international *standardization activities*, such as MPEG-4 [6] for video, JPEG-2000 [34] for still images, and MPEG-7 [24], MPEG-21 [35], SMIL [36] for generic multimedia, deal with aspects related to audiovisual (a/v) content and metadata coding and representation, aiming to provide a framework for uniformity and interoperability between developed systems. Still, mainly due to the fact that digital archives have pre-existed the developed standards, very few of them fully comply with them. In most cases, multimedia archives operate using proprietary data structures as well as administrator and end user software. Thus, the integration of multimedia archives through a common, unified access point for end users, always considering their particular copyright and access policies, emerges as a necessary step for the preservation of their content and their financial viability.

In order to achieve this goal, several research activities are currently active in the direction of *knowledge acquisition and modeling*, capturing knowledge from raw information and multimedia content in distributed repositories to turn poorly structured information into machine-processable knowledge [18][21]. A second future direction is *knowledge sharing and use*, combining semantically enriched information with context to provide infrencing for decision support and collaborative use of trusted knowledge between organizations [12]. Finally, in the *intelligent content* vision, multimedia objects integrate content with metadata and intelligence and learn to interact with devices and networks [16].

It is becoming apparent in all the above research fields that integration of diverse, heterogeneous and distributed – pre-existing – multimedia content will only be feasible through the design of *mediator systems*. In [7] for instance, a multimedia mediator is designed to provide a well-structured and controlled gateway to multimedia systems, focusing on schemas for semi-structured multimedia items and object-oriented concepts, while [3] focuses on security requirements of such mediated information systems. On the other hand, [9] deals with media abstraction and heterogeneous reasoning through the use of a unified query language for manually generated annotation, again without dealing with content or annotation semantics. A semantically rich retrieval model is suggested in [13], based on fuzzy set theory with domain-specific methods for document analysis and allowing natural language queries. Finally, [10] focuses on the design of a single intuitive interface supporting visual query languages to access distributed multimedia databases.

In this paper we present a multimedia integration approach that is focused on two parallel layers. On the *architectural layer*, the integration of heterogeneous systems is carried out using an extended 3-tier architecture for the development of the mediator system, while principles are borrowed from Message Oriented Middleware (MOM) systems [22][25], in order to increase stability and scalability. On the *content layer*, a semantic modeling approach is followed to handle problems related to the utilization of different classification schemes, different depth of detail, or different languages for the annotation of multimedia documents in individual archives. The proposed approach been utilized in the framework of the FAETHON project [4][5], where a common point of access was provided for five archives from 3 different countries.

Most existing approaches in designing multimedia mediator systems focus on specific aspects of multimedia content integration, dealing with communications protocols, metadata / annotation format, semantics, or user interface integration. The originality of the proposed system lies in the holistic approach followed to integrate individual multimedia archives both at architectural (e.g., communication, data format, metadata field mappings, MPEG-7 compliance) and content (e.g., description, classification, semantics) levels, thus affording an end-to-end solution, while still providing for maintainability and scalability in both aspects. The focus of this paper is rather on architectural design, user interface/tools and validation platform/tools, than on implementation and algorithmic details. Specifically, given a number of annotated archives, and regardless of the structure, language and format of annotation, we present an integrated 3-tier platform for integrated access. The main novelty of the approach relies in the fact that integration is achieved at both a syntactic and a conceptual level, while the main technical contribution of this paper is the automatic thematic categorization of multimedia documents given their textual annotation and a fuzzy relational knowledge base.

The structure of the paper is as follows: In Section II we present the architecture layer of the proposed integration approach, focusing on the extended 3-tier architecture. In Section III we present the content layer of the proposed integration approach, focusing on issues related to knowledge management. In Section IV we present the methodologies followed and tools developed in order to guarantee maintainability of a system that is large and complex, as well as scalability. In Section V results are presented from the application and validation of the described approach in the framework of the FAETHON project, while conclusions are drawn in Section VI.

II. Architectural Integration

Wide acceptance of the architectural approach that is known with the term 3-tier does not originate from the field of large scale systems. It originates from modern programming languages that proposed 3-tier architecture as a way to alleviate user interface and storage system dependencies from the development of the core application; Java 2 with the Model View Controller (MVC) implementation of the 3-tier architecture was the pioneer in the shift.

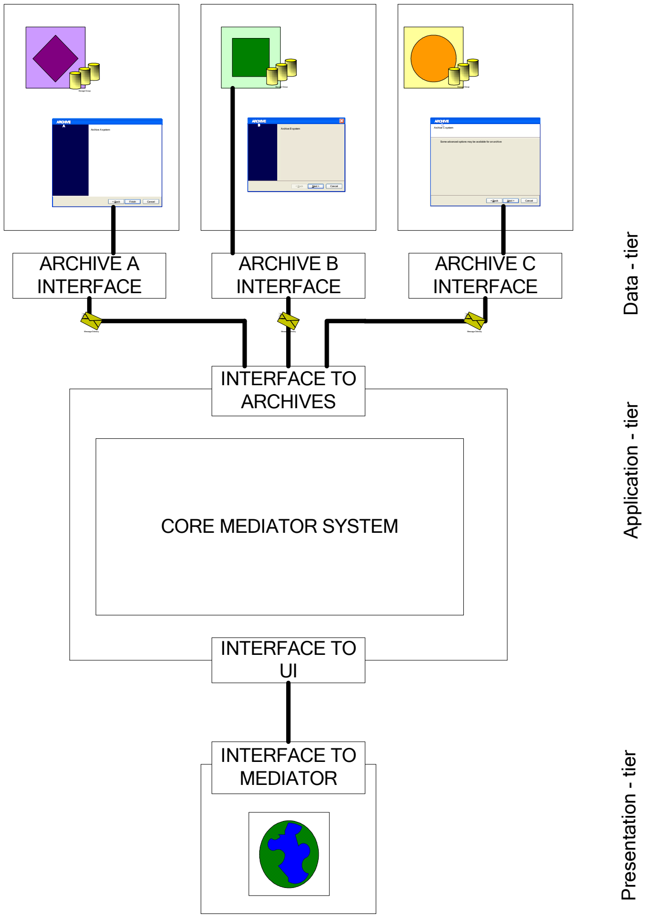


Fig. 1. The 3-tier architecture

In this framework, 3-tier refers to the definition of formal interfaces between the three main components of the system (presentation, data and application or business logic) so that they can be developed independently. In the problem examined in this work, on the other hand, the data end of the overall system, i.e. the individual multimedia archiving systems, has been developed prior to the definition of any interfacing standards. Thus, the data-tier in our approach is not implemented by the archives themselves; it is rather implemented by custom modules that allow for the interfacing of the archive systems with the newly developed mediator system.

Moreover, 3-tier architecture has been utilized mainly in applications where the 3 tiers are developed within the same software project and are running on the same system; communication between the different tiers is most typically a matter of simple function call. In the problem discussed herein the data ends of the overall system are entirely distributed and are typically located in different countries. Thus, special issues need to be considered such as the stability of the overall system in cases of network or service unavailability. An approach that follows the typical 3-tier architecture and utilizes, for example, remote procedure calls (RPC) to connect to the distributed archives would make the overall system as susceptible to errors as the sum of susceptibility of all participating entities. In this work we follow an approach inspired from MOM queued asynchronous exchange of messages that guarantees that stability of the independent systems does not affect the stability of the overall integrated system, as no assumptions are made regarding the availability of the distributed services.

The general architecture is provided in Fig. 1, where emphasis is given in the 3-tier approach followed. The presentation and application tiers are separated, as in the classic implementation of the 3-tier architecture, by defining formal interface points between them. For the individual multimedia archives, this cannot be the case. Due to the fact that they are pre-existing, it is not possible to demand for them to comply to a specific formal interfacing structure. On the other hand, failure to define such a generic interfacing scheme would result in the need to have archive specific code in the application tier. This not only goes against the whole idea of n-tier approaches which emphasizes on the independence of the tiers, but also makes it extremely difficult for the system to scale, as more archive specific code would have to be included in the application tier for each new participating archive.

In order to overcome this, the data-tier is equipped with archive interface modules. These are distributed modules residing at the site of each individual archive and act as mediators between the central system and the archive. Their role is to provide a formal interface for the core system to access the individual archives. Of course, fully automated on-line transformation of any archiving format to the MPEG-7 standard is an extremely challenging problem that remains open; what is actually supported is the identification of selected annotation features and their representation using the MPEG-7 structures [14].

In the following we provide more details on the interfaces between the different tiers, as well as on the operation of the archive interface modules.

A. Presentation - Application Tier Interface

Given that both tiers are fully developed as a result of the attempt to integrate different multimedia archives under one mediator system, the presentation and the application tiers could be developed as different sections of the same software project, with interfacing between them happening using function calls. Still, this would lessen the degree of independence between the two tiers, making the maintainability and expandability of the system questionable. Moreover, independent development may permit an open interface to third party end user interface developers, enabling alternative business scenarios.

Thus, a fully independent development approach is followed, in which XML schemas formalize the exchange of information. In the definition and development of these schemas, an important issue to consider is the nature of queries that an end user may issue as well as the nature of responses that the core system may provide for them. When accessing individual multimedia archives directly, users typically have the option to perform a search based on the structured annotation of multimedia documents; some (few) archives also allow the issuing of simple free text queries. In order for the overall integrated system to be able to offer at least as much functionality as the independent components it comprises, a very flexible communication structure needs to be defined, which will allow for the specification of possibly extremely detailed multimedia annotation information.

Such a structure has already been provided through the standardization activities of the MPEG-7 group in the development of the MPEG-7 standard. Thus, the MPEG-7 standard has been used as a basis for the definition of the user query schema; extensions are made to provide for features of information retrieval systems that have not been considered in the development of the MPEG-7 standard, as is for example the specification of relevance feedback. In the following the general structure of the UserAction Description Scheme is provided; its components are further decomposed in the actual schemas of the system.

```
<?xml version="1.0" encoding="UTF-8"?>
<!-- edited with XML Spy v4.3 U (http://www.xmlspy.com)
by Manolis Wallace (Image, Video and Multimedia
Laboratory) -->
<xs:schema targetNamespace="urn:FAETHON:schema:2001"
xmlns:faethon="urn:FAETHON:schema:2001"
xmlns:xs="http://www.w3.org/2001/XMLSchema"
xmlns:mpeg7="urn:mpeg:mpeg7:schema:2001"
elementFormDefault="qualified"
attributeFormDefault="unqualified">
    <xs:import namespace="urn:mpeg:mpeg7:schema:2001"
schemaLocation="./mpeg/MPEG-7-2001.xsd"/>
    <xs:import namespace="urn:FAETHON:schema:2001"
schemaLocation="./FAETHON-2001.xsd"/>
    <xs:complexType name="UserAction">
        <xs:complexContent>
            <xs:extension base="faethon:Request">
                <xs:choice>
                    <xs:element name="userQuery"
type="faethon:UserQuery"/>
                    <xs:element name="relevanceFeedBack"/>
                    <xs:element name="browse"/>
                    <xs:element name="view"/>
                </xs:choice>
            </xs:extension>
        </xs:complexContent>
    </xs:complexType>
    <xs:complexType name="FaethonResult">
        <xs:complexContent>
            <xs:restriction base="faethon:Result">
```

```
                <xs:sequence>
                    <xs:element name="faethonResponse"
type="mpeg7:Mpeg7Type"/>
                </xs:sequence>
            </xs:restriction>
        </xs:complexContent>
    </xs:complexType>
    <xs:complexType name="UserQuery">
        <xs:sequence>
            <xs:element name="KeywordQuery"
type="mpeg7:KeywordAnnotationType" minOccurs="0"/>
            <xs:element name="Metadata"
type="faethon:FaethonMetaDataType" minOccurs="0"/>
        </xs:sequence>
    </xs:complexType>
    <xs:complexType name="RelevanceFeedback"/>
    <xs:complexType name="Browse"/>
    <xs:complexType name="View"/>
</xs:schema>
```

Similarly, the responses of the core system (the application tier) are formalized through XML schemas. Once again, the nature of the responses that may be provided have been considered in the development of the schemas. Since the option to request for annotation details has been allowed in the specification of the UserAction type, MPEG-7 structural Description Schemes (DSs) are utilized in order to allow for the transmission of structured multimedia annotation between the two tiers.

```
<?xml version="1.0" encoding="UTF-8"?>
<!-- edited with XML Spy v4.2 U (http://www.xmlspy.com)
by Manolis Wallace (Image, Video and Multimedia
Laboratory) -->
<schema targetNamespace="urn:FAETHON:schema:2001"
xmlns="http://www.w3.org/2001/XMLSchema"
xmlns:faethon="urn:FAETHON:schema:2001"
xmlns:mpeg7="urn:mpeg:mpeg7:schema:2001"
elementFormDefault="qualified"
attributeFormDefault="unqualified">
    <import namespace="urn:mpeg:mpeg7:schema:2001"
schemaLocation="./mpeg/MPEG-7-2001.xsd"/>
    <complexType name="SystemEvent" abstract="true"/>
    <complexType name="Request" abstract="true">
        <complexContent>
            <extension base="faethon:SystemEvent"/>
        </complexContent>
    </complexType>
    <complexType name="Reply">
        <complexContent>
            <extension base="faethon:SystemEvent">
                <choice minOccurs="0"
maxOccurs="unbounded">
                    <element name="result"
type="faethon:Result"/>
                    <element name="error"
type="faethon:Error"/>
                </choice>
            </extension>
        </complexContent>
    </complexType>
    <complexType name="Result" abstract="true">
        <sequence>
            <any namespace="##targetNamespace"
minOccurs="0" maxOccurs="unbounded"/>
        </sequence>
    </complexType>
    <complexType name="Error">
        <sequence>
            <element name="message" type="string"
minOccurs="0" maxOccurs="unbounded"/>
        </sequence>
    </complexType>
    <!-- FaethonMetaDataType -->
    <!-- The list of metadata is just an example. Note
that not all fields of the corresponding
    MPEG-7 DS may be relevant.-->
    <complexType name="FaethonMetaDataType">
        <sequence>
            <element name="MediaIdentification"
type="mpeg7:MediaIdentificationType" minOccurs="0"/>
            <element name="MediaFormat"
type="mpeg7:MediaFormatType" minOccurs="0"/>
            <element name="MediaQuality"
type="mpeg7:MediaQualityType" minOccurs="0"/>
            <element name="Creation"
type="mpeg7:CreationType" minOccurs="0"/>
            <element name="Classification"
type="mpeg7:ClassificationType" minOccurs="0"/>
        </sequence>
    </complexType>
</schema>
```

In the provided schema only a few of the structural DSs provided by MPEG-7 are included; this list of DSs is enough to provide for the specification of all the annotation elements that have been selected in the user requirements and system specifications phase to be supported by the system. Alternatively, all MPEG-7 structural DSs could be included in the definition. Still, although the theoretical merits of such an approach may be obvious, there would be no practical benefit as currently there are no user interfaces that are able to present all the range of annotation details supported by MPEG-7; a choice of elements needs to be made for any real life system.

The network layer of the communication is built using Web Services over HTTP. Alternative, SSL technology can be easily used to replace the clear text transmission method of HTTP in order to provide for enhanced security features, if required, without affecting the overall operation and communication of the tiers.

B. *Application - Data Tier Interface*

Communication between the core mediator system and the individual archives is performed in two stages; the archive interface has the role of the mediator in this process. Thus, the core system communicates with the archive interface and the archive interface with the individual archive.

As far as the former is concerned, an approach quite similar to that of the communication between the presentation and the application tiers is followed. Thus, system requests are defined using a schema that utilizes MPEG-7 DSs in order to provide the option of specifying multimedia structural details in the request.

```
<?xml version="1.0" encoding="UTF-8"?>
<!-- edited with XML Spy v4.3 U (http://www.xmlspy.com)
by Manolis Wallace (Image, Video and Multimedia
Laboratory) -->
<xs:schema targetNamespace="urn:FAETHON:schema:2001"
xmlns:xs="http://www.w3.org/2001/XMLSchema"
xmlns:faethon="urn:FAETHON:schema:2001"
xmlns:mpeg7="urn:mpeg:mpeg7:schema:2001"
```

```
elementFormDefault="qualified"
attributeFormDefault="unqualified">
   <xs:import namespace="urn:FAETHON:schema:2001"
schemaLocation=".\FAETHON-2001.xsd"/>
   <xs:complexType name="FaethonQuery">
      <xs:sequence>
         <xs:element name="document" type="xs:anyURI"
minOccurs="0" maxOccurs="unbounded"/>
         <xs:element name="metaData"
type="faethon:FaethonMetaDataType" minOccurs="0"
maxOccurs="unbounded"/>
      </xs:sequence>
      <xs:attribute name="getNewDocuments"
type="xs:boolean" default="false"/>
   </xs:complexType>
   <xs:complexType name="SemanticExpression"/>
   <xs:complexType name="ArchiveResponse">
      <xs:complexContent>
         <xs:restriction base="faethon:Result">
            <xs:sequence minOccurs="0"
maxOccurs="unbounded">
               <xs:element name="weight"
type="mpeg7:zeroToOneType"/>
               <xs:sequence>
                  <xs:element name="document">
                     <xs:complexType>
                        <xs:complexContent>
                           <xs:extension
base="mpeg7:Mpeg7Type"/>
                        </xs:complexContent>
                     </xs:complexType>
                  </xs:element>
               </xs:sequence>
            </xs:sequence>
         </xs:restriction>
      </xs:complexContent>
   </xs:complexType>
</xs:schema>
```

The replies of the archive interface module are similarly formalized using an XML schema that borrows from MPEG-7. This is in fact the same Reply Description Scheme that is utilized for exchange of replies between the presentation and the application tiers.

The main difference between the presentation-application interface and the application-data interface is that for the development of the latter special care needs to be taken in order to ensure system stability. The blocking nature of Web Services communication – the entity submitting a request is forced to wait for the response before continuing – is prohibitive for the interface of the core system with the individual archives, as the unavailability of one of the archives should not be reason enough to disrupt the normal operation of the overall system; the property of graceful degradation is a required characteristic for a large scale integration project.

Thus, instead of web services, MOMs are used as the backbone carrier of information between the core system and the archive interfaces. The main characteristic of MOMs that makes them suitable for the task is that they are designed for asynchronous exchange of messages. In other words, communication is non blocking – the calling procedure simply specifies the data it would like to have, but does not stop and wait for it; when that data is made available it is processed. As far as the low level communication channel is concerned, that can be provided by any technology. Practically, the result is that the core system appears to be available regardless of the availability of the individual archives that participate; when one of them is not available, data from the remaining archives only may be processed and considered in the preparation of the system's response.

As far as the communication between the archive interface and the actual archive is concerned, no formal communication structure is specified. This is in fact a custom software component that is developed separately for each archive in order to provide for its interfacing with the overall system. As can be seen in Fig. 1, the archive interface may reside at the location of the archive, but it is not a part of its system. The reason for this is that the normal operation of the archive should not be upset by the time consuming (and in some cases questionable and unstable) process of system integration. Moreover, any other communication modules that are connected to the archive should be allowed to continue to operate normally, as the maintenance of all existing applications is of great financial importance for multimedia archives.

The archive interface communicates with the archive database management system using the same protocols as the existing archive proprietary end user interface, as is the case with archive B of Fig. 1. Requests that may be issued by the core system are "translated" based on this communication protocol, and the archive response is also translated to the MPEG-7 format described by the system schemas. Of course, any intelligence that has been built in the archive's proprietary end user interface is lost and has to be re-coded in the archive interface. In cases, though, where the proprietary end user application of the archive allows for software interfaces to connect to it (archives A and C in Fig. 1) the archive interface utilizes the software in order to search in the archive, thus taking advantage of its intelligence.

Overall, the archive interface modules undertake one of the most important parts of the overall system integration process, as they are in charge of the translation of proprietary annotation available in the archive to standard MPEG-7 annotation. The cost of multimedia annotation is such – only expert users can perform it and it is an extremely time consuming process – that without a software module to do this automatically, it would not be financially feasible for a non compliant archive to re-annotate its content and start participating in such an integrated access system. Moreover, exactly because of the cost of the annotation, content holders would be extremely reluctant to participate in an integrated system that disregards parts of the existing annotation.

C. Core Mediator

The structure of the core mediator system, presented as a black box in Fig. 1, is presented to a greater detail in Fig. 2. The interface to archives at the top and the interface to UI at the bottom have also been presented in the previous subsections – the interaction components are in charge of accepting data and forwarding to the core system while the presentation components are in charge of accumulating data from the core system, formatting it and forwarding to the communications component.

The system operates in two distinct and parallel modes: query and update. In *query mode*, the query analysis module is the recipient of the semantic part of the user query, while the search engine accepts the metadata part. The semantic part is specified by textual keywords and is matched to indexing information already stored in the DBMS of the core system, while the metadata part is specified through the MPEG-7 descriptors and is forwarded to the participating archives. Responses of those archives that respond before a timer threshold expires are merged.

Of course, in order for the query analysis module to be able to provide a response, an appropriate indexing of documents needs to have already been performed. This is performed in *offline, update mode* of operation by the DTC and DECO modules. It is worth mentioning that the offline mode is not a mode of operation that requires the system to be unavailable; the “offline” term simply refers to the fact that there is no strict time frame within which it is imperative for the participating processes to have provided their output, and thus more sophisticated algorithms may be utilized. In fact the DTC and DECO modules are always active, indexing new documents or refreshing indexing information based on newly acquired knowledge, much like web crawlers of internet search engines that continuously crawl the web indexing new pages or refreshing the indexing information for known ones. In the same mode, accumulated usage information can be processed for the update of the user profiles.

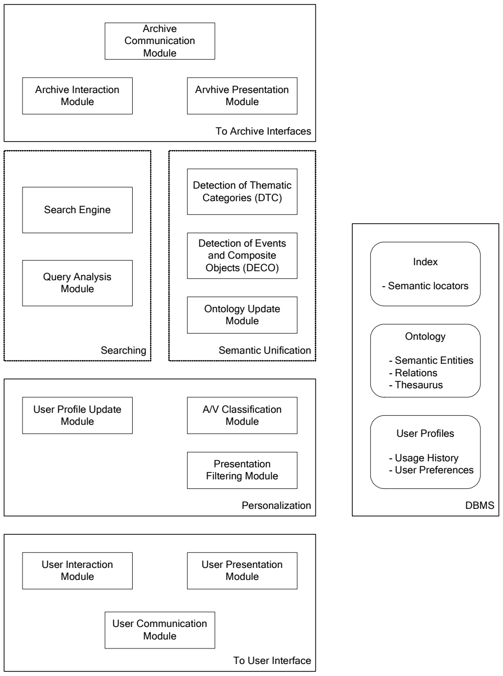


Fig. 2. The structure of the core mediator system

Other system components include the index, ontology and ontology update module discussed in section III, and the personalization subsystem along with the user profiles, which are outside the scope of this paper and discussed in [8], [29] and [32].

III. Content Integration

Section II proposes an approach that allows for the systemic integration of diverse – in architectural nature – multimedia archives under the umbrella of a mediator system for search and access. Still, the issue of integration with respect to content organization and access also needs to be handled before the system is able to operate in a truly integrated manner. For example, suppose that two archives with annotations in different languages are integrated using the above-mentioned approach. Queries to both can be made through a common interface and results will be provided and presented in a uniform manner. Still, it will not be possible to have results from both archives at the same time – which was the initial goal behind the integration procedure – as only one of the two archives will have documents with annotation matching the query, depending on the language in which the user issued the query. The same is especially true for multimedia items with no or little associated textual annotation, which have to be semantically analyzed and indexed if matching to textual queries is to be made possible.

In order to cope with different languages, different levels of detail in annotation, choice of different terms for the description of similar situations and objects and so on, a semantic modeling approach is followed. In this approach, indexing of multimedia documents in the core system is not based on keywords but rather on semantic entities; each semantic entity can have multiple definitions, possibly in even different languages. A query can be posed in a language, e.g., in English, and returned results may be originally annotated in another, e.g. German or Greek. Moreover, the utilization of semantic entities in indexing enables us to utilize ontological information in order to provide enhanced searching and profiling services.

A. *Knowledge Model*

Similarly to having to re-annotate content, it is not feasible, financially and time wise, to perform the semantic indexing manually. Thus, an automated process for the matching of document annotations to semantic entities needs to be specified. Therefore, the definitions of the semantic entities need to be available to the system; they are located in the DBMS, in the section labeled as Ontology in Fig. 2. In order to handle the problem of different forms of terms, alternatively to the solution of stemming, we may choose to include all possible forms of the terms in the definition of entities. Using a similar approach, we may include terms in various languages, thus allowing for uniform handling of multimedia documents annotated using different languages.

A different problem that needs to be solved, ideally handled with Natural Language Processing (NLP) techniques, is that of terms found in textual annotation that match more than one semantic entity definitions. As the problem of NLP remains open, a simpler, yet efficient, approach is followed to determine the correct mapping of terms found in the annotation to the set of known entities: in the field of ontologies the exact meaning of a word is determined by the "context" [17][23]. Following the same approach, we utilize an ontological knowledge representation and use the context to determine the correct way to map the terms to semantic entities.

An ontological description contains (at least) a set of entity definitions together with the taxonomic relations among them. The two combined form a *semantic encyclopedia*. Using the entity definitions we may determine the set of candidate entities for each term, whereas using the taxonomic relations we may define and estimate the context of a sentence or short paragraph, thus solving the problem of term to entity mapping. The main elements of the schema used to describe the semantic encyclopedia are included in Fig. 3, Fig. 4, Fig. 5 and Fig. 6.

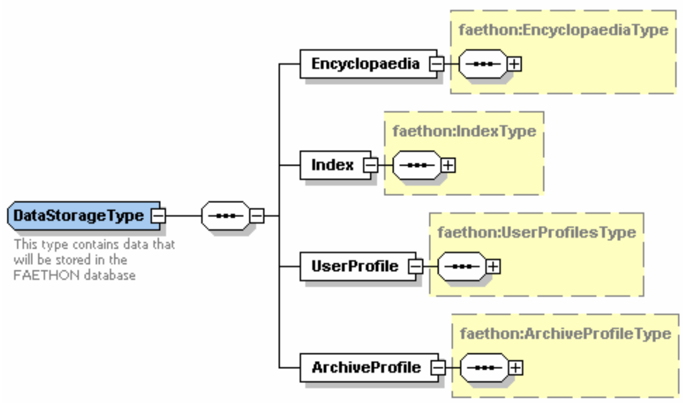

faethon:EncyclopaediaType
Encyclopaedia
faethon:IndexType
Index
DataStorageType
This type contains data that will be stored in the FAETHON database
faethon:UserProfilesType
UserProfile
faethon:ArchiveProfileType
ArchiveProfile

Fig. 3. Graphical representation of DataStorageType

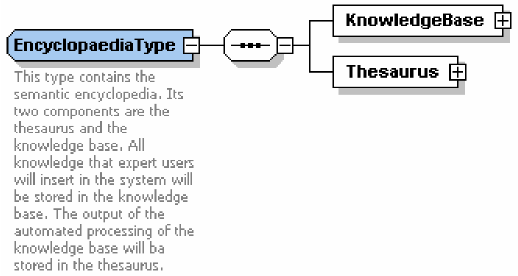

EncyclopaediaType
KnowledgeBase
Thesaurus
This type contains the semantic encyclopedia. Its two components are the thesaurus and the knowledge base. All knowledge that expert users will insert in the system will be stored in the knowledge base. The output of the automated processing of the knowledge base will ba stored in the thesaurus.

Fig. 4. Graphical representation of EncyclopaediaType

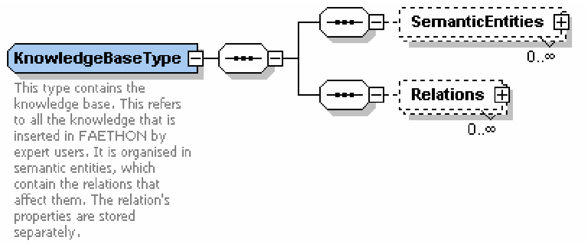

KnowledgeBaseType
SemanticEntities
0..∞
Relations
0..∞
This type contains the knowledge base. This refers to all the knowledge that is inserted in FAETHON by expert users. It is organised in semantic entities, which contain the relations that affect them. The relation's properties are stored separately.

Fig. 5. Graphical representation of KnowledgeBaseType

A semantic encyclopedia, or equivalently an ontological knowledge description, in order to be highly descriptive, needs to contain a large number of distinct and diverse – in their semantic nature – relations among semantic entities. As a result, available information will be divided among them, making each one of them inadequate to fully describe a semantic context. Thus, more than one such relations may need to be combined in order to provide a view of the knowledge that suffices for context definition and estimation.

For the purpose of analyzing multimedia document descriptions we use a view that has been generated with the use of the following fuzzy semantic taxonomic relations [31], whose semantics are defined in MPEG-7:

- Part *P*, inverted. *P*(*a*,*b*)>0 indicates that *b* is an a part of *a*. For example, *a* may be "human body" and *b* may be "hand".
- Specialization *Sp*. *Sp*(*a*,*b*)>0 indicates that *b* is a specialization of *a*. For example, *a* may be "vehicle" and *b* may be "car".
- Example *Ex*. *Ex*(*a*,*b*)>0 indicates that *b* is an example of *a*. For example, *a* may be "player" and *b* may be "Jordan".
- Instrument *Ins*. *Ins*(*a*,*b*)>0 indicates that *b* is an instrument of *a*. For example, *a* may be "music" and *b* may be "drums".
- Location *Loc*, inverted. *L*(*a*,*b*)>0 indicates that *b* is the location of *a*. For example, *a* may be "concert" and *b* may be "stage".
- Patient *Pat*. *Pat*(*a*,*b*)>0 indicates that *b* is a patient of *a*. For example, *a* may be "course" and *b* may be "student".
- Property *Pr*, inverted. *Pr*(*a*,*b*)>0 indicates that *b* is a property of *a*. For example, *a* may be "Jordan" and *b* may be "star".

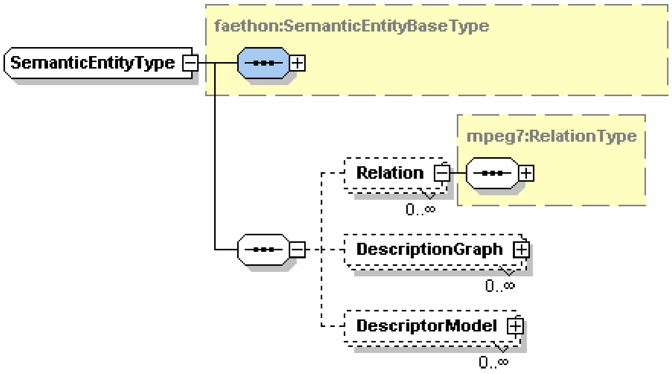


Fig. 6. Graphical representation of SemanticEntityType.

Thus, the utilized view *T* is a new semantic relation that is calculated as:

$$T = \left(Sp \cup P^{-1} \cup Ins \cup \Pr^{-1} \cup Pat \cup Loc^{-1} \cup Ex\right)^{(n-1)} \quad (1)$$

The (*n-1*) exponent indicates *n-1* compositions, which are guaranteed to establish the property of transitivity for the view [15]; it is necessary to have the view in a closed transitive form, in order to be able to answer to questions such as "*which entities are related to entity x?*" in *O*(log*n*) instead of *O*(n^2) times, where *n* is the (very large) count of known semantic entities.

Given the fact that the complexity for a single fuzzy relation composition is *O*(n^3), as well as the fact that maintenance of the encyclopedia may call for frequent changes in the abovementioned relations, it is ease to see that this leads to a bottle neck. For this purpose, a more efficient, computationally wise, methodology that is targeted especially to sparse relations of the form of the relations contained in the semantic encyclopedia has been developed [33]. This algorithm establishes transitivity with a complexity that is below $O(n^2)$ and updates the transitive relation in the case of an addition of an element with a sub-linear complexity, thus making the overall approach tractable.

Based on the semantics of the participating relations, it is easy to see that T is ideal for the determination of the topics that an entity may be related to, as well as for the estimation of the common meaning, i.e. the context, of a set of entities.

B. The Notion of Context

When using an ontological description, it is the context of a term that provides its truly intended meaning. In other words, the true source of information is the co-occurrence of certain entities and not each one independently. Thus, in the process of multimedia indexing we have to use the common meaning of terms in order to best determine the entities to which they should be mapped. We refer to this as their *context*; in general, term *context* refers to whatever is common among a set of elements. Relation T will be used for the detection the context of a set of semantic entities [2].

Relying on the semantics of relation T, we define the context $K(s)$ of a single semantic entity $s \in S$ as the set of its antecedents in relation T, where S is the set of all semantic entities contained in the encyclopedia. More formally, $K(s) = T_{\leq}(s)$, following the standard superset – subset notation from fuzzy relational algebra. We demand that, when A is a normal fuzzy set, the "considered" context $\mathcal{K}(s)$ of s, i.e. the entity's context when taking its degree of participation to the set into account, is low when the degree of participation $A(s)$ is high, or when the context of the crisp entity $K(s)$ is low. Therefore

$$cp(\mathcal{K}(s))=cp(K(s))\cap(S \cdot A(s)) \qquad (2)$$

where cp is an involutive fuzzy complement. By applying de Morgan's law, we obtain:

$$\mathcal{K}(s)=K(s)\cup cp(S \cdot A(s)) \qquad (3)$$

Then the overall context of the set is again easily calculated as

$$K(A) = \bigcap \mathcal{K}(s_i), s_i \in A \qquad (4)$$

Considering the semantics of the T relation and the proposed process of context determination, it is easy to realize that when the entities in a set are highly related to a common meaning, the context will have high degrees of membership for the entities that represent this common meaning. Therefore, the height of the context $h(K(A))$, i.e. the greatest membership degree that appears in it, may be used as a measure of the semantic correlation of entities in set A. We will refer to this measure as *intensity* of the context.

C. Context-Sensitive Interpretation – DECO module

The definitions of semantic entities in the semantic encyclopedia contain sequences of labels, each one providing a different textual form of the semantic entity, possibly in more than one languages. Matching those to the terms in the annotation of a multimedia document (or a specific segment of a multimedia document) we can acquire the semantic representation of the annotation.

Interpretation needs to take place simultaneously with context estimation. We follow the following method: Let a phrase or a paragraph of the textual annotation contain the terms $\{t_i\}$ with $i = 1,\ldots,T$. Let also t_i be the textual description of semantic entities $\{s_{ij}\}$ with $j = 1,\ldots, T_i$. Then, there exist $N_A = \prod_i T_i$ distinct combinations of semantic entities that may be used for the representation of the annotation; for each one of those we calculate the corresponding context.

The intensity of the context is a semantic measure of the association of the entities in a set. Thus, out of the candidate semantic interpretations $\{a_k\}$, where $k = 1,2,...,N_A$, the one that produces the most intense context is the one that contains the semantic entities that are most related to each other; this is the combination that is chosen as output of the process of semantic annotation interpretation.

$$a = a_i \in \{a_1,...,a_{N_A}\} : h(a_i) \geq h(a_j) \forall a_j \in \{a_1,...,a_{N_A}\} \qquad (5)$$

This semantic annotation interpretation is exhaustive, in the sense that it needs to consider all possible interpretations of a given annotation. Still, this is not a problem in the framework where it is applied as:

- Phrases and small paragraphs do not contain large numbers of terms
- The number of distinct semantic entities that may have a common textual description is not large.
- The gain in the quality of the semantic content of the interpreted annotation is largely more important than the added computational burden.

D. Context-Sensitive Understanding – DTC module

After the simple semantic indexing of textual annotation, a multimedia document d (or equivalently a segment of a multimedia document) is represented only by its mapping to semantic entities. More formally, documents D are indexed by semantic entities S via the semantic index I:

$$I : S \to D \qquad (6)$$

Therefore, the context of a document is again defined via the semantic entities that are related to it. This contains valuable information that can assist in the process of content analysis and understanding [27][28][30][31]. Before detecting the topics that are related to a document d, the set of semantic entities that are related to it needs to be clustered, according to their common meaning. More specifically, the set to be clustered is the support of the document:

$${}^{0+}d = \{s \in S : I(s,d) > 0\} \qquad (7)$$

Most clustering methods belong to either of two general categories, partitioning and hierarchical [26]. Partitioning methods create a crisp or fuzzy clustering of a given data set, but require the number of clusters as input. Since the number of topics that exist in a document is not known beforehand, partitioning methods are inapplicable for the task at hand; a hierarchical clustering algorithm needs to be applied [20]. Hierarchical methods are divided into agglomerative and divisive. Of those, the first are more widely studied and applied, as well as more robust.

The two key points in hierarchical clustering are the identification of the clusters to merge at each step, i.e. the definition of a meaningful measure for *CI*, and the identification of the optimal terminating step, i.e. the definition of a meaningful termination criterion.

When clustering semantic entities, the ideal association measure for two clusters c_1, c_2 is one that quantifies their semantic correlation. In the previous we have defined such a measure: the intensity of their common context $h(K(c_1 \cup c_2))$. The process of merging should terminate when the entities are clustered into sets that correspond to distinct topics. We may identify this case by the fact that no pair of clusters will exist with a common context of high intensity. Therefore, the termination criterion shall be a threshold on the *CI*.

To determine the topics that are related to a cluster c^{final}, two things need to be considered: the scalar cardinality of the cluster $|c^{final}|$ and its context. Since context has been defined only for normal fuzzy sets, we need to first normalize the cluster as follows:

$$c^{normal}(s) = \frac{c^{final}(s)}{h(c^{final}(s))}, \forall s \in {}^{0+}d \tag{8}$$

Obviously, semantic entities that are not contained in the context of c^{normal} cannot be considered as being related to the topic of the cluster. Therefore the thematic categorization R_T of the extracted clusters is

$$R_T(c^{final}) \subseteq R_T^*(c^{normal}) = w(K(c^{normal})) \tag{9}$$

where *w* is a weak modifier [15]. Modifiers, which are also met in the literature as *linguistic hedges*, are used to adjust mathematically computed values so as to match their semantically anticipated counterparts.

In the case where the semantic entities that index document *d* are all clustered in a unique cluster c^{final}, then $R_T(d)=R_T^*(c^{normal})$ is a meaningful approach. On the other hand, when multiple clusters are detected, then it is imperative that cluster cardinalities are considered as well.

Clusters of extremely low cardinality probably only contain misleading entities, e.g. entities that correspond to terms that were used in a metaphorical manner, and therefore need to be ignored in the estimation of $R_T(d)$. On the contrary, clusters of high cardinality almost certainly correspond to the distinct topics that *d* is related to, and need to be considered in the estimation of $R_T(d)$. The notion of "high cardinality" is modeled with the use of a "large" fuzzy number $L(\cdot)$; $L(a)$ is the truth value of the preposition "*a* is high", and, consequently, $L(|b|)$ is the truth value of the preposition "the cardinality of cluster *b* is high".

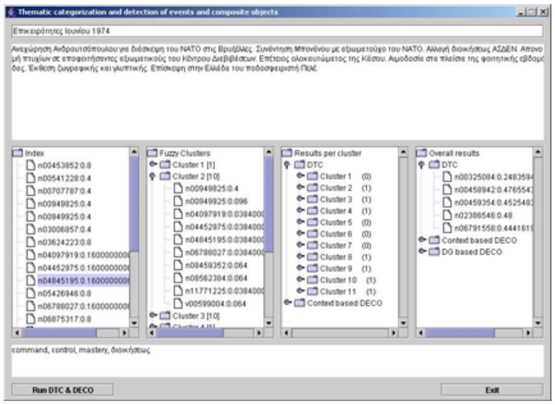
The set of topics that correspond to a document is the set of topics that correspond to each one of the detected clusters of semantic entities that index the given document.

$$R_T(d) = \bigcup R_T(c^{final}), c^{final} \in G \tag{10}$$

where $\cup$ is a fuzzy co-norm and *G* is the set of fuzzy clusters that have been detected in *d*. The topics that are related to each cluster are computed, after adjusting membership degrees according to scalar cardinalities, as follows:

$$R_T(c^{final}) = R_T^*(c^{normal}) \cdot L(c^{final}) \tag{11}$$

It is easy to see that $R_T(s,d)$ will be high if a cluster c^{final}, whose context contains *s*, is detected in *d*, and additionally, the cardinality of c^{final} is high and the degree of membership of *s* in the context of the cluster is also high (i.e., if the topic is related to the cluster and the cluster is not comprised of misleading entities).


Thematic categorization and detection of events and composite objects
Run DTC & DECO
Exit

Επικαιρότητες Ιουνίου 1974

Αναχώρηση Ανδρουτσόπουλου για διάσκεψη του ΝΑΤΟ στις Βρυξέλλες. Συνάντηση Μπονάνου με αξιωματούχο του ΝΑΤΟ. Αλλαγή διοικήσεως ΑΣΔΕΝ. Απονομή πτυχίων σε αποφοιτήσαντες αξιωματικούς του Κέντρου Διαβιβάσεων. Επέτειος ολοκαυτώματος της Κάσου. Αιμοδοσία στα πλαίσια της φοιτητικής εβδομάδας. Έκθεση ζωγραφικής και γλυπτικής. Επίσκεψη στην Ελλάδα του ποδοσφαιριστή Πελέ.

```
Index
  n00453852:0.8
  n00541228:0.4
  n00707787:0.4
  n00949825:0.4
  n00949925:0.4
  n03006857:0.4
  n03624223:0.8
  n04097919:0.160000000
  n04452975:0.160000000
  n04845195:0.160000000
  n05426946:0.8
  n06788027:0.160000000
  n06875317:0.8

Fuzzy Clusters
  Cluster 1 [1]
  Cluster 2 [10]
    n00949825:0.4
    n00949925:0.096
    n04097919:0.0384000
    n04452975:0.0384000
    n04845195:0.0384000
    n06788027:0.0384000
    n08459352:0.064
    n08562384:0.064
    n11771225:0.0384000
    v00599004:0.064
  Cluster 3 [10]
  Cluster 4 [1]

Results per cluster
  DTC
    Cluster 1    (0)
    Cluster 2    (1)
    Cluster 3    (1)
    Cluster 4    (1)
    Cluster 5    (0)
    Cluster 6    (0)
    Cluster 7    (0)
    Cluster 8    (1)
    Cluster 9    (1)
    Cluster 10   (1)
    Cluster 11   (1)
  Context based DECO

Overall results
  DTC
    n00325084:0.2483594
    n00458942:0.4765541
    n00459354:0.4525483
    n02386546:0.48
    n06791558:0.4441619
  Context based DECO
  DG based DECO
```

command, control, mastery, διοικήσεως

Fig. 7. The document analysis tool

The combined DECO and DTC procedures have been initially developed and tested independently of the other core system components. The original document analysis tool is presented in Fig. 3. In a similar manner, documents in the usage history of a user can be analyzed for the specification of the user profile [29][32].

IV. Maintenance and Scalability

In order for a large integrated system such as the one described herein, in which administratively independent archives participate, the addition of new archives and removal of other archives cannot be considered a rare situation. Quite the contrary, depending on financial opportunities and strategic alliances, the list of participating archives may vary frequently. Therefore, in order for the system to be viable, these processes need to be thoroughly prescribed and assurance is needed that their execution will not upset the normal operation of the overall integrated system. Moreover, as the goal of the system is to integrate as many archives as possible, scalability issues also need to be considered. In this section we discuss these two issues.

A. Maintenance Issues

Maintainability refers to the ability to cope with small, but possibly frequent, changes in the requirements or environment settings of operation of a system. In the context of the integrated system described herein, maintainability mainly refers to the ability to integrate new archives and update/expand the knowledge stored in the semantic encyclopedia.

As far as the incorporation of new archives is concerned, this is a relatively time consuming, yet fully prescribed procedure:

- The archive interface has to be developed and linked to the search and indexing processes.
- In the case of an archive that contains terms and descriptions in some language other than English translations of terms have to be added to the knowledge base.
- In the case that the archive content refers to topics not already found in archives attached to the FAETHON system, the available knowledge may have to be extended with more semantic relation entries as to optimize the output of the semantic processes.
- Semantic indexing needs to be performed for the content of the new archive by the DECO and DTC processes.

What is primarily required is the development of a custom archive interface, which will allow for the proprietary annotation utilized by the archive to be translated to the MPEG-7 compliant representation scheme supported by the core mediator system. The operation of this software module can be further decomposed in the following tasks:

- Parsing the FaethonQuery requests.
- Executing the corresponding queries on the archive. Depending on the archive this could be a query posed to an existing archival service, to a database or a search in a raw data file.
- Encoding the response using the formal ArchiveResponse format and returns the resulting XML file to the core mediator system.

FaethonQuery and ArchiveResponse are formal DSs extending the MPEG-7 standard defined ones, similar to the ones presented earlier in the paper.

In cases where the archive's annotation scheme is either very poor, or already MPEG-7 compliant, this becomes a trivial procedure. In cases, though, where the archive is following a complex, custom annotation scheme, every element of this scheme needs to be modeled to the corresponding MPEG-7 Description Scheme, so that absolutely no information is lost in the process of the integration.

Moreover, the communication schemas used between the application-tier and data-tier of the system are made available to interested multimedia archives. In this way, multimedia archives that do not already have a fully digitized storage and annotation scheme, as well as archives that wish to upgrade/replace their existing system, can develop a scheme that is compliant with the overall integrated scheme. Such archives can be integrated in the system directly, without the need for an archive interface. Finally, the same archiving software as the one used by two of the already participating archives can be modified and customized for any interested archive. This system comes with valuable supporting software, as is the powerful annotator tool presented in Fig. 4 [1][11][14]. In that case, the already developed archive interface can be utilized.

The most challenging part of the maintenance of the system is related to the knowledge stored in the semantic encyclopedia. This needs to be edited mainly in two cases: when a new language needs to be supported in user queries or when a new archive is integrated in the system. In the former case, the task is currently virtually intractable, as all defined semantic entities need to be manually edited – the system currently holds definitions for approximately 110000 semantic entities. Advances in the WordNet project are expected to provide a solution to this problem [19]; as WordNet was the original source for the generation of the semantic entities which can now be used by end users – in English – to specify their semantic queries, a mapping between synsets of WordNet for different languages can be automatically merged into the semantic encyclopedia.

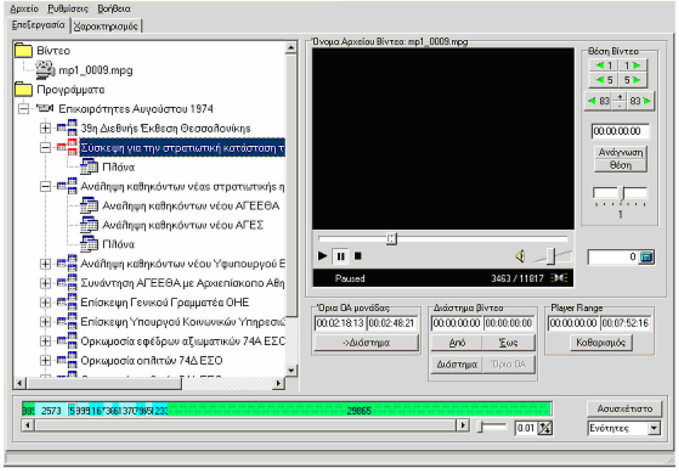
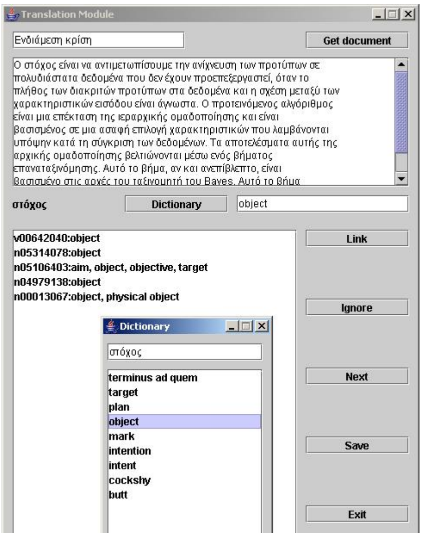

Αρχείο
Ρυθμίσεις
Βοήθεια
Επεξεργασία
Χαρακτηρισμός
Βίντεο
mp1_0009.mpg
Προγράμματα
Επικαιρότητες Αυγούστου 1974
39η Διεθνής Έκθεση Θεσσαλονίκης
Σύσκεψη για την στρατιωτική κατάσταση τ
Πλάνα
Ανάληψη καθηκόντων νέας στρατιωτικής η
Αναληψη καθηκόντων νέου ΑΓΕΕΘΑ
Ανάληψη καθηκόντων νέου ΑΓΕΣ
Πλάνα
Ανάληψη καθηκόντων νέου Υφυπουργού Ε
Συνάντηση ΑΓΕΕΘΑ με Αρχιεπίσκοπο Αθη
Επίσκεψη Γενικού Γραμματέα ΟΗΕ
Επίσκεψη Υπουργού Κοινωνικών Υπηρεσιώ
Ορκωμοσία εφέδρων αξιωματικών 74Α ΕΣΟ
Ορκωμοσία οπλιτών 74Δ ΕΣΟ
Όνομα Αρχείου Βίντεο: mp1_0009.mpg
Θέση Βίντεο
1
1
5
5
83
83
00:00:00.00
Ανάγνωση Θέση
1
0
Paused
3463 / 11817
Όρια ΘΑ μονάδας
00:02:18:13
00:02:48:21
->Διάστημα
Διάστημα βίντεο
00:00:00.00
00:00:00:00
Από
Έως
Διάστημα
Όρια ΘΑ
Player Range
00:00:00.00
00:07:52:16
Καθαρισμός
2573
29865
0.01
Ασυσχέτιστο
Ενότητες

Fig. 8. The annotator application

In the case of new archives joining the integrated system, small adjustments may need to be made to the semantic entity definitions or to the ontological taxonomies. If the archive is not annotated using the English language, terms utilized in the annotation may not be mapped to any of the semantic entities in the semantic encyclopedia. The DECO module detects this and alerts the system administrator.

A translation tool has been developed in order to facilitate the semantic encyclopedia with the missing terms. The tool, which is presented in Fig. 5, parses an annotation text and identifies the terms that are not already found in the encyclopedia. The administrator (or the expert user in general) can then select the semantic entity (or entities) to which the term should be added. In order to further facilitate this process, the tool is integrated with a dictionary, allowing the user to specify the corresponding English term; the system presents all the semantic entities that contain the term in their definition and the user can select the one(s) to which the new term belongs.

Once an archive has been integrated through the development of the archive interface and the extension of the encyclopedia in order to contain the new terms, documents can be indexed by DECO and DTC, and thus made available to the end users through semantic and metadata searches. Still, in most cases the newly encountered documents do not appear frequently in the result sets. The reason is that if their indexing does not match the query terms exactly they cannot be identified as relevant to it if the knowledge in the encyclopedia does not describe this fact; the taxonomical part of the encyclopedia has been developed manually in a trial and error manner, and thus this is not a rare case. In order for the relations in the encyclopedia to be properly updated, a suitable end user interface has been developed; using this interface the expert user can identify which archives should be considered, and which should not, when performing a search. Using this option the expert user can identify missing elements in the encyclopedia by issuing queries and studying the results – when documents are missing from the results the relational knowledge needs to be extended. This, of course, requires that the expert user is also aware of the actual content of the new archive.



Ενδιάμεση κρίση

Ο στόχος είναι να αντιμετωπίσουμε την ανίχνευση των προτύπων σε
πολυδιάστατα δεδομένα που δεν έχουν προεπεξεργαστεί, όταν το
πλήθος των διακριτών προτύπων στα δεδομένα και η σχέση μεταξύ των
χαρακτηριστικών εισόδου είναι άγνωστα. Ο προτεινόμενος αλγόριθμος
είναι μια επέκταση της ιεραρχικής ομαδοποίησης και είναι
βασισμένος σε μια ασαφή επιλογή χαρακτηριστικών που λαμβάνονται
υπόψην κατά τη σύγκριση των δεδομένων. Τα αποτελέσματα αυτής της
αρχικής ομαδοποίησης βελτιώνονται μέσω ενός βήματος
επαναταξινόμησης. Αυτό το βήμα, αν και ανεπίβλεπτο, είναι
βασισμένο στις αρχές του ταξινομητή του Bayes. Αυτό το βήμα

στόχος object

v00642040:object
n05314078:object
n05106403:aim, object, objective, target
n04979138:object
n00013067:object, physical object

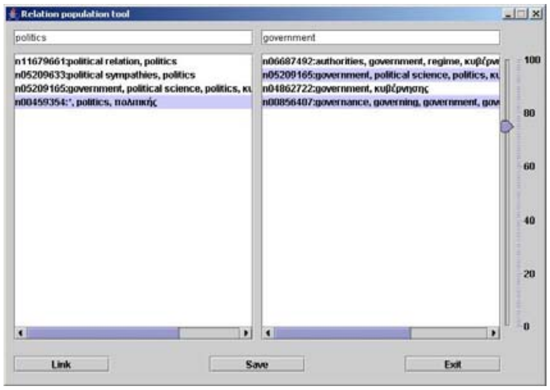
στόχος

terminus ad quem
target
plan
object
mark
intention
intent
cockshy
butt

Fig. 9. The translation tool

For the extension of the fuzzy relations in the encyclopedia the tool presented in Fig. 6 is used. Using this tool the expert user can identify source and destination semantic entities by specifying corresponding words, as well as the degree to which these semantic entities are to be linked. It implements the algorithm introduced in [33], in order to re-establish transitivity of the relation in sub-linear time. At the end of the updating procedure for any of the relations in the encyclopedia, the *T* relation is also recalculated, again using a computationally efficient approach [33].

In the cases that the knowledge is extended to a great extend – for example with the definition and population of new semantic relations – the definition of the semantic view *T* might also need to be changed. For more complex knowledge management operations as this one the more sophisticated knowledge management tool presented in Fig. 7 is utilized.



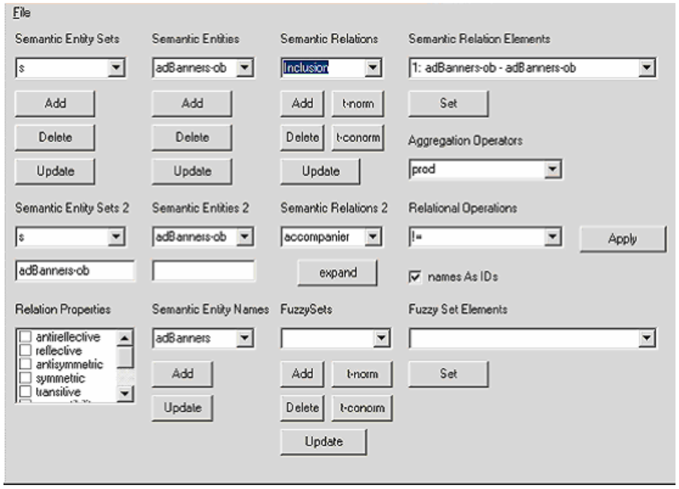
politics

n11679661:political relation, politics
n05209633:political sympathies, politics
n05209165:government, political science, politics, κυ
n00459354:', politics, πολιτικής

government

n06687492:authorities, government, regime, κυβέρνη
n05209165:government, political science, politics, κυ
n04862722:government, κυβέρνησης
n00856407:governance, governing, government, gov

Fig. 10. The relation population tool



Semantic Entity Sets: s
Semantic Entities: adBanners-ob
Semantic Relations: Inclusion
Semantic Relation Elements: 1: adBanners-ob - adBanners-ob
Aggregation Operators: prod
Semantic Entity Sets 2: s
adBanners-ob
Semantic Entities 2: adBanners-ob
Semantic Relations 2: accompanier
Relational Operations: !=
☑ names As IDs
Relation Properties: ☐ antireflective ☐ reflective ☐ antisymmetric ☐ symmetric ☐ transitive
Semantic Entity Names: adBanners
FuzzySets
Fuzzy Set Elements

Fig. 11. The knowledge management tool

B. *Scalability Issues*

The presented system operates as a mediator between users and multimedia archives. Therefore, scalability can be defined with respect either to the number of users accessing the system or to the number of participating archives.

As far as the former is concerned, the issue is trivial. To the end user, the proposed system is a Web Interface that provides feedback based on information stored either locally in the DBMS (for semantic keyword based queries) or in the distributed participating archives (for metadata queries). In either case, the existence of multiple users does not impose an important overhead for the system. Similar web applications with numerous users have existed for years (for example Amazon and Google) and experience shows that as long as the bandwidth of the line connecting the server to the backbone through which most users access it is sufficient, the number of concurrent users does not create any problem. This is especially true as, using indexing techniques, processing of user queries has been optimized as analyzed below.

As far as the count of participating archives is concerned, once again the scaling of the system is excellent. When it comes to the generation of the semantic index, scaling is linear, as each document needs to be accessed once in order for DECO and DTC modules to process it. Thus, participation of more archives can easily be handled in this sense with the initiation of multiple indexing processes. Moreover, indexing services run in offline mode, and thus the overhead in their operation does not truly upset the operation of the system.

When, on the other hand it comes to the execution of user queries, two cases need to be considered. In the case of a metadata query, the query is forwarded by the Archive Communication Module (as shown in Fig. 3) to all participating archives in parallel, and thus the existence of multiple archives does not augment the overall time. The probability for any archive to be unavailable is of course larger, but as we have already explained the MOM approach followed does not allow this to affect the operation of the system. Finally, in the case of semantic queries, the list of matching documents is readily available in the semantic index. With more participating archives, it is expected – assuming that all archives have roughly the same size – that the size of the semantic index will increase linearly. Still, due to the ordered representation model of the semantic index in memory in the form of a semantic relation [33], this results in only a logarithmic raise in the overall processing time. In other words, 1000 times more archives need to be inserted in the system in order for the time to process a user query to be augmented by a factor of 10, or, equivalently, in order for 10 times more processing power to be needed to produce an equally rapid response.

Overall, the proposed integration architecture and approach guarantees excellent scaling properties in all senses. In any case, under extreme conditions, the system has been designed to enable operation using clusters of servers to process queries in a parallel fashion. However, such conditions would be very far from the actual conditions under which the system was tested (up to five archives and 10,000 concurrent users).

V. Integrated System

In this section we provide information acquired from the practical application of the integration methodology and system described in this work in the framework of the EU funded FAETHON project [4]. We start by describing the architecture and content of the five multimedia archives that have been integrated, and continue by listing findings from the integration and validation procedures.

A. Participating Archives

The five archives that have been integrated in the framework of the FAETHON project are ERT (Hellenic Broadcasting Corporation), FAG (Film Archive Greece), FAA (Film Archive Austria), ORF (Austrian Broadcasting Corporation) and Alinari Archive. Three different archive interfaces have been implemented for the these archives, as ERT and FAG use similar archiving systems and are integrated in the overall FAETHON system using identical archive interfaces; the same goes for FAA and ORF.

The archive system of ERT is based on a centralized software solution, the PANORAMA audiovisual content management application system, utilizing the MPEG-7 content description standard and ORACLE relational DBMS technology. The archive consists of two servers, namely (i) the database server, which hosts the archive database with all a/v content metadata, and (ii) the media server / web server / web service provider, which hosts all the a/v content itself, handles media streaming, provides an end-user web interface to the archive content. Annotation in the ERT archive (as well as in the FAG archive) is in Greek. The ERT archive interface is responsible for the communication between the FAETHON system and the ERT archive system. It parses and translates requests made by the core mediator system to the proprietary XML format utilized by the ERT system. The PANORAMA Search API is then activated to search the ERT database. Finally, the response produced by the API is assembled into an MPEG-7 compliant format and returned to FAETHON. The implementation and interface of the FAG archive has been similar to that of ERT, with the exception of several custom metadata field mappings and classification schemes.

Annotation in the FAA archive (as well as in the ORF archive) is in German. The FAA archive interface provides functionality for querying the publicly available data of FAA by SOAP via HTTP. The queries accepted by that interface are translated into the native format of FAA’s system. The interface also includes a database access layer which performs the connection to FAA’s ORACLE database. The interface is implemented in Java and acts on top of a TOMCAT servlet container, while SOAP requests and responses are handled by a WASP server. A layered approach with partially generic components was followed in the design and implementation of the FAA archive interface, which allows for adaptation based on the needs of other local systems. Due to this feature, it was possible to adapt the developed archive interface in order to also support the integration of the ORF archive into the system.

The Alinari archive consists of a large set of still images, along with associated metadata including, for example, title, description, date, photographer, artist, etc., stored in an MS-SQL Server database. Annotation in the Alinari archive is in English and is considerably less detailed than the annotation in the other participating archives. The Alinari archive interface is responsible for the communication between the FAETHON system and the Alinari archive system. It parses the core system request in order to extract the requirements forming a query to directly communicate with the Alinari database. The response produced by the database is assembled into an MPEG-7 compliant format and returned to FAETHON. The Alinari archive interface has also been implemented in Java allowing utilization on different platforms.

B. Addition of New Archives

In the first prototype, only ERT and FAA archives were integrated into the system. ORF, FAG and the Alinari archive have been integrated into the system additionally to the originally participating archives, in order to increase user involvement in FAETHON and validate its demonstrator. This process has also provided the opportunity to develop and test the prescribed procedure for the addition of new archives.

In the case of the FAG and ORF archives, the integration process was rather trivial. As far as the systemic integration is concerned, these archives use archival systems similar to those of the originally participating archives, and thus it was not necessary to develop new custom archive interfaces. As far as the semantic integration is concerned, their content is of similar nature to the content of FAA and ERT and their annotation languages the same. Thus, full integration at both architecture and content layer was completed easily with only minor maintenance overhead for the translation of some new terms and the addition of some new relational elements in the semantic encyclopedia.

In the case of the Alinari archive, on the other hand, a complete integration procedure had to be completed. The new archive interface was developed in a rather timely manner; during the development of the archive interface both the FAETHON system and the Alinari Archive operated uninterrupted by the preparations for the forthcoming integration. Due to the formal specification of the interfacing data structures between the core mediator system and the archive interface, the systemic integration of Alinari into the core system once the archive interface was developed was automatic – there was no fine tuning required.

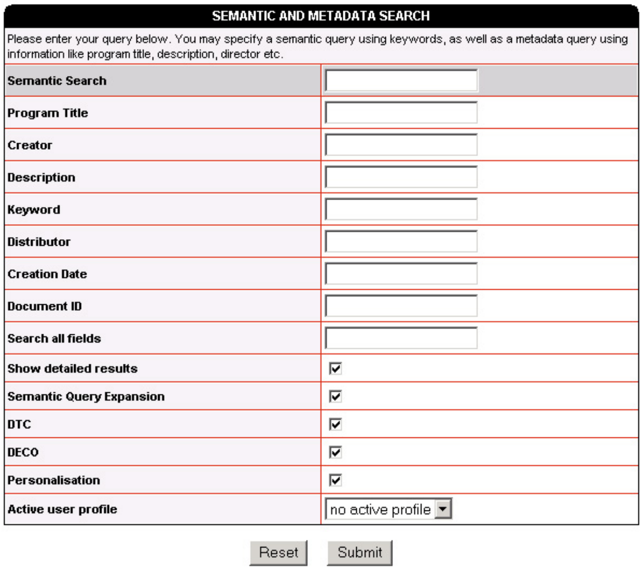
As the annotation language for Alinari is English, there was no need for term translation – the core system has been equipped with a mapping of all English terms to their corresponding semantic entities through processing of the WordNet synsets. On the other hand, due to the fact that the archived content was radically different in nature than that of the other participating archives, excessive updating of the semantic relations in the semantic encyclopedia was required. This process was also completed successfully, as is made evident by the fact that the overall integrated system is now able to correctly respond to queries in a seamless manner including related multimedia documents from all archives in its response.

C. *Verification Procedures*

Based on the orientation and nature of the software developed for the intelligent modules of the FAETHON system, the functional, or black-box testing methodology, for integrated system testing has been used. That is, validation and verification is based on the specification to develop test cases, while the code is invisible. The next step is integration testing, to ensure that all subsystems and components have been integrated correctly and there are no faults in their interfacing. Performance testing follows, to verify that the performance of the integrated system in terms e.g. of capacity, response time or throughput is accepted. Finally, scalability testing is employed to investigate how the system behaves under increased workload environments. Depending on the nature of each specific test, a combination of the user interface depicted in Fig. 8 and Fig. 9 and automated scripts was used to access the system and log its operation.

The purpose of *functional verification* is to verify that the integrated system performs according to its user requirements and system specifications. The main findings and conclusions drawn from the functional verification of the FAETHON system are outlined here:

1. The search engine supports degrees of relevance and all metadata included in the user query form. Semantic indexing is done without communication with the archives, while no archive data are transferred to the FAETHON system without the archive's agreement. The result of using the FAETHON system clearly outperforms the union of the results of using the individual archives.
2. The DBMS fully supports the FAETHON Schema, and its query language is rich enough to support the operation of the system. The encyclopedia can describe the semantic content of a/v documents and can be manually updated and automatically expanded, checked for consistency and corrected.
3. The user interface supports all metadata accessed by user queries. A very small part of MPEG-7 descriptions is shown graphically, while the entire descriptions can be viewed or downloaded in XML. Downloading of multimedia content is supported.
4. In the FAETHON system, the knowledge stored in the encyclopedia has been both acquired by public sources and specifically developed for the participating archive content; it includes a sufficient number of semantic entities to index all multimedia documents available. All modules are integrated to the extent that a query can be processed.

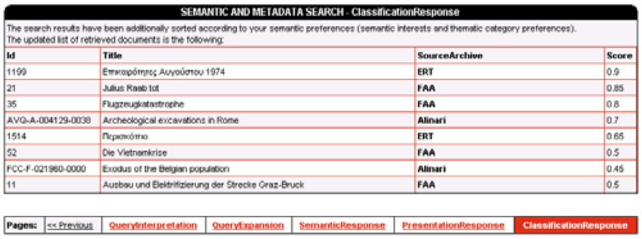


SEMANTIC AND METADATA SEARCH

Please enter your query below. You may specify a semantic query using keywords, as well as a metadata query using information like program title, description, director etc.

Semantic Search	
Program Title	
Creator	
Description	
Keyword	
Distributor	
Creation Date	
Document ID	
Search all fields	
Show detailed results	☑
Semantic Query Expansion	☑
DTC	☑
DECO	☑
Personalisation	☑
Active user profile	no active profile

Fig. 12. The query section of the end user interface used for testing



SEMANTIC AND METADATA SEARCH - ClassificationResponse

The search results have been additionally sorted according to your semantic preferences (semantic interests and thematic category preferences).
The updated list of retrieved documents is the following:

Id	Title	SourceArchive	Score
1199	Επικαιρότητες Αυγούστου 1974	ERT	0.9
21	Julius Raab tot	FAA	0.85
35	Flugzeugkatastrophe	FAA	0.8
AVQ-A-004129-0038	Archeological excavations in Rome	Alinari	0.7
1514	Περισκόπιο	ERT	0.65
52	Die Vietnamkrise	FAA	0.5
FCC-F-021960-0000	Exodus of the Belgian population	Alinari	0.45
11	Ausbau und Elektrifizierung der Strecke Graz-Bruck	FAA	0.5

Fig. 13. The results' page of the end user interface; documents from different archives are seamlessly incorporated

Integration tests are aimed to the detection and correction of faults due to interface errors or invalid assumptions about interfaces. Integration errors include interface misuse, interface misunderstanding and timing errors. The main findings are outlined here:

1. All interfaces between FAETHON components developed by different organizations, thus requiring integration, were of the message passing type, i.e., subsystems request services from other subsystems by an appropriate message in XML format according to the FAETHON Schemata. Full conformance to these schemata was verified, and thus no correcting actions were needed.
2. These message passing interfaces have been (i) between the central FAETHON server and the archive interfaces, using the FaethonQuery and ArchiveResponse types, (ii) between the central FAETHON server and the end user interface, using the UserQuery and Faethon Response type, and (iii) inside the central FAETHON server for integration of the independent intelligent modules.
3. All FaethonQuery options have been tested by manually generating combinations of all possible Faethon queries to archive interfaces, even with conflicting options (e.g. requesting all documents and documents after a specific date, at the same time). Tests have been successful for all archive interfaces and only minor modifications were needed.
4. No interface misuse or timing error was observed in any interface. Minor interface misunderstandings occurred around the schema of the FaethonQuery and ArchiveResponse types, which were fixed and the schema was updated when needed.
5. Overall, the use of XML data exchange and the strict specification of the FAETHON Schemata was proven to have significantly simplified the integration procedure and minimized to a considerable extent any integration errors.

Performance testing and verification involve configuring and tuning the infrastructure of the system, both hardware and software, until it can support the number of concurrent users with an acceptable response time. The purpose is to detect bottlenecks of the system and tune the software so as to increase system speed, capacity and efficiency, and to shift or reduce the load by employing alternative data representations or caching techniques. The main findings are the following:

1. Construction of the semantic index from all participating archives using the DTC and DECO process is an offline process performed when the archive content or the knowledge is updated and can performed within a few minutes for the archives currently participating. Most of this time is dedicated to the transfer of data between the distributed archives and the core mediator system, while the processing time for DECO and DTC is considerably smaller.
2. In online mode of operation, the processes of query analysis, lookup in the semantic index by the search engine, a/v classification and presentation filtering only require a few milliseconds, despite the size of the FAETHON knowledge base (tens of thousands of terms, semantic entities and relations). Any delays are therefore caused by network traffic or internal processing of XML archive responses.
3. Network traffic turns out to cause no serious performance problem. Testing over extended periods of time has shown that network delays during processing of a query require again a few milliseconds. Archives are contacted by the core system in a parallel fashion, and when a connection fails or is timed out, limited cached information is used instead.

Based on the above, the only real bottleneck of the system proved to be internal processing of XML archive responses, whose size in certain cases tend to be in the order of Megabytes. Mining of information in these XML documents, marshalling and unmarshalling in several steps for the purpose of communication between internal intelligent modules some times required a processing time in the order of seconds, making it hard to support a large number (e.g. larger than 10) concurrent users. This bottleneck was fixed by employing alternative internal data representations, and caching techniques for all information produced internally between intermediate processing steps. The work load was considerably reduced and system speed and efficiency accordingly increased. The final prototype can on average serve a user query within a few milliseconds.

Scalability testing includes examination of the system behavior under conditions of increased number of users and workload (transaction size), keeping the size of the database and number of connected archives fixed, as the FAETHON system already includes the total of all five participating archives a/v content and metadata available. Experiments included load, stress and capacity testing in order to determine (i) whether the system can sustain, in a stable and reliable way, a specific number of users with acceptable response times, and (ii) the maximum number of concurrent users and transactions the system can support. The findings and conclusions are given below:

1. The five archives participating in the integrated system contain in total around 3,000 a/v documents whose complete metadata description in XML takes up around 10 Megabytes. The knowledge base, on the other hand, contains around 110,000 semantic entities mapped to around 160,000 terms in three languages (mainly English),

and around 80,000 relation elements.

2. Keeping the above parameters fixed, a large number of coordinated users was emulated by software scripts and consistent test repetitions were carried out, automatically measuring system response times. It was found that the integrated system can serve on average around 300 transactions (queries) per second, keeping the response time below 2 seconds. As more users are connected to the system, the above figure slightly decreases. The test was carried out with up to 20 concurrent connected virtual users, where this figure decreased to around 250 transactions per second on average. The decrease is due the processing of additional user profiles during the response personalization phase.
3. The above result included queries according to random combinations of terms from the knowledge base, along with random metadata queries. Since processing times for metadata queries largely depend on archive and archive interface response sizes, separate experiments were carried out to estimate the deviations of the overall system's processing rate. It was found that the integrated system can serve from around 1-2 (for the largest possible response, containing around 1,000 documents) to around 1,200 (for a response of a single document) transactions per second, again keeping the response time below 2 seconds.
4. Load and stress testing was carried out and it was verified that the system can sustain the above transaction rates with acceptable response times (always below 5 seconds) over extended periods of time, around 2 days of continuous operation at maximum, in a stable and reliable way, with no errors.
5. Capacity testing was carried out by configuring run-time settings to define the way that the scripts run, in order to accurately emulate real user behavior. The settings included think time depending on user experience (varying between 5 and 60 seconds) and connection speed (varying between 30 and 10,000 Kbps). In this experiment, it was found the integrated system can support around 4,000 concurrent connected users with average response times of 2 seconds.

VI. Conclusion

In this paper, we have addressed the problem of integration of distributed multimedia archives under the umbrella of a common access mediator. This is a complicated integration problem that has to be tackled on two different levels. In the architectural level, interfaces between heterogeneous systems need to be specified, while in the content level semantic integration needs to be pursued.

As far as the architectural integration is concerned, we have followed an extended 3-tier approach with influences from Messaging Oriented Middleware. Though the specification and development of archive interfaces as independent components the normal operation of the integrated archives is not disturbed by the integration process, while the messaging approach to data exchange increases the stability and robustness of the overall system. As far as the content integration is concerned, semantic indexing through automatic document analysis has been employed. This has allowed the seamless participation of archives with annotations in different languages and to different detail in the same retrieval system.

The described approach has been utilized to integrate five European audiovisual archives from three different countries. The procedure required in order for a new archive to be added to the integrated system has also been prescribed, supported with maintenance tools and tested successfully in three different cases. Functional, integration, performance and scaling verification tests have indicated that the proposed integration approach is both stable and efficient.

Our future research orientations will include the development of methods and tools aimed at higher levels of information representation, extraction and use, including automated knowledge discovery, metadata / annotation extraction, summarization, conceptual and contextual retrieval of multimedia content, emphasizing on cross-media and cross-lingual aspects.

Acknowledgment

The knowledge management tool presented in Fig. 7 has been developed by Dr. Giorgos Akrivas as part of his work in the FAETHON project and his PhD at the Image, Video and Multimedia Systems Laboratory of NTUA under the supervision of Professor Stefanos Kollias.

This work has been partially funded by the HERACLETUS project of the Greek Ministry of Education.

References

[1] Akrivas G., Ioannou S., Karakoulakis E., Karpouzis K., Avrithis Y., Delopoulos A., Kollias S., Varlamis I. and Vaziriannis M., "An Intelligent System for Retrieval and Mining of Audiovisual Material Based on the MPEG-7 Description Schemes", Eunite 2001, European Symposium on Intelligent Technologies, Hybrid Systems and their implementation on Smart Adaptive Systems, Tenerife, Spain

[2] Akrivas, G., Wallace, M., Andreou, G., Stamou, G. and Kollias, S., "Context-Sensitive Semantic Query Expansion", Proceedings of the IEEE International Conference on Artificial Intelligence Systems (ICAIS), Divnomorskoe, Russia, September 2002

[3] Altenschmidt C., Biskup J., Flegel U. and Karabulut Y., "Secure mediation: requirements, design, and architecture", Journal of Computer Security, Volume 11, Issue 3, March 2003, pp. 365 – 398.

[4] Athanasiadis T., Avrithis Y., "Adding Semantics to Audiovisual Content: The FAETHON Project", Proc. of International Conference for Image and Video Retrieval (CIVR '04), Dublin, Ireland, July 21-23, 2004

[5] Avrithis Y., Stamou G., Wallace M., Marques F., Salembier P., Giro X., Haas W., Vallant H., Zufferey M., "Unified Access to Heterogeneous Audiovisual Archives", Journal of Universal Computer Science 9(6), pp. 510-519, 2003

[6] Battista S., Casalino F., Lande C., "MPEG-4: A Multimedia Standard for the Third Millenium", Part 2. IEEE Multimedia 7 (1), pages 76-84, 2000

[7] Biskup J., Freitag J., Karabulut Y., and Sprick B., "A mediator for multimedia systems", Proceedings 3rd International Workshop on Multimedia Information Systems, Como, Italy, Sept. 1997.

[8] Bloehdorn S., Simou N., Tzouvaras V., Petridis K., Handschuh S., Avrithis Y., Kompatsiaris I., Staab S. and Strintzis M.G., "Knowledge Representation for Semantic Multimedia Content Analysis and Reasoning", in Proc. of European Workshop on the Integration of

Knowledge, Semantics and Digital Media Technology (EWIMT 04), Lon-don, U.K., November 25-26, 2004.

[9] Brink A., Marcus S., Subrahmanian V., "Heterogeneous Multimedia Reasoning", IEEE Computer, Volume 28 , Issue 9, Sept. 1995, pp. 33-39.

[10] Cruz I., James K., "A user-centered interface for querying distributed multimedia databases", Proceedings of the 1999 ACM SIGMOD International Conference on Management of Data, Philadelphia, PA, 1999, pp. 590-592.

[11] Daras P., Kompatsiaris I., Grinias I., Akrivas G., Tziritas G., Kollias S. and Strintzis M., "MPEG-4 Authoring Tool Using Moving Object Segmentation and Tracking in Video Shots", EURASIP Journal of Applied Signal Processing (JASP), volume 2003, issue 9, pages 861-877

[12] Euzenat J., Fensel D., Lara R. and Gómez-Pérez A., "Knowledge Web: Realising The Semantic Web, all The Way to Knowledge Enhanced Multimedia Documents", Proc. of European Workshop on the Integration of Knowledge, Semantics and Digital Media Technology (EWIMT), London, U.K., November 25-26, 2004.

[13] Glöckner I., Knoll A., "Natural Language Navigation in Multimedia Archives: An Integrated Approach", Proceedings of the 7th ACM International Conference on Multimedia, Orlando, FL, pp. 313-322.

[14] Ioannou S., Akrivas G., "Archiving multimedia content descriptions: an early adaptation of MPEG-7", Proceedings of Int. Workshop on Very Low Bitrate Video Coding (VLBV01)

[15] Klir G. and Bo Yuan, Fuzzy Sets and Fuzzy Logic, Theory and Applications, (New Jersey, 1995) Prentice Hall.

[16] Kompatsiaris I., Avrithis Y., Hobson P. and Strinzis M.G., "Integrating Knowledge, Semantics and Content for User-Centred Intelligent Media Services: the aceMedia Project", Proc. of Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS), Lisboa, Portugal, April 21-23, 2004.

[17] Maedche A., Motik B., Silva N. and Volz R., MAFRA - An Ontology Mapping FRAmework in the Context of the SemanticWeb. Proceedings of the Workshop on Ontology Transformation at ECAI2002, Lyon, France, July 2002

[18] Mich O., Brunelli R. and Modena C.M., "A survey on video indexing", Journal of Visual Communications and Image Representation, Volume 10, 1999, pp. 78-112.

[19] Miller G.A., Ed. "WordNet: An on-line lexical database", International Journal of Lexicography, Vol. 3, Issue 4, Winter 1990, pp. 235-312.

[20] Miyamoto S., Fuzzy Sets in Information Retrieval and Cluster Analysis, Kluwer Academic Publishers, Dordrecht / Boston / London, 1990

[21] Naphade M.R., Kozintsev I.V., and Huang T.S., "A factor graph framework for semantic video indexing", IEEE Trans. on Circuits and Systems for Video Technology, Volume 12, Issue 1, Jan. 2002, pp. 40-52.

[22] Rao B.R., "Making the Most of Middleware", Data Communications International 24, 12 (September 1995):89-96

[23] Schnurr H.-P., Erdmann M., Maedche A.and Staab S., „From Manual to Semiautomatic Semantic Annotation: About Ontology-based Text Annotation Tools", in Proc. COLING 2000 Workshop on Semantic Annotation and Intelligent Content, Luxembourg, August 5-6, 2000.

[24] Sikora T.., "The MPEG-7 Visual standard for content description - an overview", IEEE Trans. on Circuits and Systems for Video Technology, special issue on MPEG-7, 11(6):696-702, June 2001

[25] Steinke S., "Middleware Meets the Network", LAN: The Network Solutions Magazine 10 ,13 (December 1995):56

[26] Theodoridis S. and Koutroumbas K., Pattern Recognition, Academic Press, 1998

[27] Wallace M., Akrivas G., Mylonas P., Avrithis Y., and Kollias S., "Using Context and Fuzzy Relations to Interpret Multimedia Content," In Proc. 3rd International Workshop on Content-Based Multimedia Indexing, CBMI'03, Rennes, France, September 22-24 2003

[28] Wallace M., Akrivas G., Stamou G., "Automatic Thematic Categorization of Documents Using a Fuzzy Taxonomy and Fuzzy Hierarchical Clustering", Proceedings of the IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2003), St. Liouis, MO, USA, May 2003

[29] Wallace M., Akrivas G., Stamou G. and Kollias S., "Representation of user preferences and adaptation to context in multimedia content --based retrieval", Proceedings of the Workshop on Multimedia Semantics, SOFSEM 2002: Theory and Practice of Informatics, Milovy, Czech Republic, November 2002

[30] Wallace M., Athanasiadis T., Avrithis Y., Stamou G., Kollias S., "A mediator system for hetero-lingual audiovisual content", Proceedings of the International Conference on Multi-platform e-Publishing, November 2004, Athens, Greece

[31] Wallace M., Avrithis Y., Stamou G., Kollias S., "Knowledge-based Multimedia Content Indexing and Retrieval", Stamou G., Kollias S. (Editors), Multimedia Content and Semantic Web: Methods, Standards and Tools, Wiley, 2004

[32] Wallace M. and Stamou G., "Towards a Context Aware Mining of User Interests for Consumption of Multimedia Documents", Proceedings of the IEEE International Conference on Multimedia (ICME), Lausanne, Switzerland, August 2002

[33] Wallace M., Avrithis Y., Kollias S., "Computationally efficient sup-t transitive closure for sparse fuzzy binary relations", Fuzzy Sets and Systems, in press

[34] ISO/IEC JTC1/SC29/WG1 N1646R: JPEG 2000 Part I Final Committee Draft Version 1.0, 2000

[35] MPEG-21 Overview v.3, ISO/IEC JTC1/SC29/WG11 N4511, December 2001

[36] Synchronized Multimedia Integration Language (SMIL). Recommendation REC-smil-1998 0615, World Wide Web Consortium (W3D), 1998. http://www.w3.org/TR/REC-smil.